



## RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/858,332E  
Source: IFW/6  
Date Processed by STIC: 8/26/04

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER** **VERSION 4.2 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box-1450, Alexandria, VA 22313-1450
3. Hand-Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 06/05/04):  
U.S. Patent and Trademark Office, 220 20<sup>th</sup> Street S., Customer Window, Mail Stop Sequence, Crystal Plaza Two, Lobby, Room 1B03, Arlington, VA 22202

Revised 05/17/04



IFW16

## RAW SEQUENCE LISTING

DATE: 08/26/2004

PATENT APPLICATION: US/09/858,332E

TIME: 11:22:51

Input Set : D:\SEQUENCE LISTING.TXT

Output Set: N:\CRF4\08262004\I858332E.raw

4 <110> APPLICANT: Tchaga, Grigory S.  
 5       Jokhadze, George  
 7 <120> TITLE OF INVENTION: Metal Ion Affinity Tags and Methods for  
 8       Using the Same  
 11 <130> FILE REFERENCE: CLON-056CIP  
 13 <140> CURRENT APPLICATION NUMBER: US 09/858,332E  
 14 <141> CURRENT FILING DATE: 2001-05-15  
 16 <150> PRIOR APPLICATION NUMBER: 09/404,017  
 17 <151> PRIOR FILING DATE: 1999-09-23  
 19 <150> PRIOR APPLICATION NUMBER: 60/101,867  
 20 <151> PRIOR FILING DATE: 1998-09-25  
 22 <160> NUMBER OF SEQ ID NOS: 27  
 24 <170> SOFTWARE: FastSEQ for Windows Version 4.0  
 26 <210> SEQ ID NO: 1  
 27 <211> LENGTH: 16  
 28 <212> TYPE: PRT  
 29 <213> ORGANISM: Artificial Sequence  
 31 <220> FEATURE:  
 32 <223> OTHER INFORMATION: affinity peptide  
 34 <400> SEQUENCE: 1  
 35 His Leu Ile His Asn Val His Lys Glu Glu His Ala His Ala His Asn  
 36   1                   5                   10                   15  
 39 <210> SEQ ID NO: 2  
 40 <211> LENGTH: 18  
 41 <212> TYPE: PRT  
 42 <213> ORGANISM: Artificial Sequence  
 44 <220> FEATURE:  
 45 <223> OTHER INFORMATION: affinity peptide  
 47 <400> SEQUENCE: 2  
 48 His Asp Asp His Asp Asp His Asp Asp His Asp Asp His Asp Asp His  
 49   1                   5                   10                   15  
 50 Asp Asp  
 54 <210> SEQ ID NO: 3  
 55 <211> LENGTH: 18  
 56 <212> TYPE: PRT  
 57 <213> ORGANISM: Artificial Sequence  
 59 <220> FEATURE:  
 60 <223> OTHER INFORMATION: affinity peptide  
 62 <400> SEQUENCE: 3  
 63 His Glu Glu His Glu Glu His Glu Glu His Glu Glu His  
 64   1                   5                   10                   15  
 65 Glu Glu  
 69 <210> SEQ ID NO: 4

Does Not Comply  
 Corrected Diskette Needed

pp 2,6

## RAW SEQUENCE LISTING

DATE: 08/26/2004

PATENT APPLICATION: US/09/858,332E

TIME: 11:22:51

Input Set : D:\SEQUENCE LISTING.TXT

Output Set: N:\CRF4\08262004\I858332E.raw

70 <211> LENGTH: 18  
 71 <212> TYPE: PRT  
 72 <213> ORGANISM: Artificial Sequence  
 74 <220> FEATURE:  
 75 <223> OTHER INFORMATION: affinity peptide  
 77 <400> SEQUENCE: 4  
 78 His Asp Glu His Asp Glu His Glu Asn His Glu Asn His Glu Asp His  
 79 1 5 10 15  
 80 Glu Asp  
 84 <210> SEQ ID NO: 5  
 85 <211> LENGTH: 18  
 86 <212> TYPE: PRT  
 87 <213> ORGANISM: Artificial Sequence  
 89 <220> FEATURE:  
 90 <223> OTHER INFORMATION: affinity peptide  
 92 <400> SEQUENCE: 5  
 93 His Glu Asp His Glu Asp His Glu Asp His Glu Asp His  
 94 1 5 10 15  
 95 Glu Asp  
 99 <210> SEQ ID NO: 6  
 100 <211> LENGTH: 5  
 101 <212> TYPE: PRT  
 102 <213> ORGANISM: Artificial Sequence  
 104 <220> FEATURE:  
 106 <223> OTHER INFORMATION: *exploration mandatory in <2207-2237 section*  
 106 <400> 6  
 107 Asp Asp Asp Asp Lys  
 108 1 5  
 111 <210> SEQ ID NO: 7  
 112 <211> LENGTH: 4  
 113 <212> TYPE: PRT  
 114 <213> ORGANISM: Artificial Sequence  
 116 <220> FEATURE:  
 117 <223> OTHER INFORMATION: enterokinase cleavage site  
 119 <400> SEQUENCE: 7  
 120 Ile Glu Gly Arg  
 121 1  
 124 <210> SEQ ID NO: 8  
 125 <211> LENGTH: 6  
 126 <212> TYPE: PRT  
 127 <213> ORGANISM: Artificial Sequence  
 129 <220> FEATURE:  
 130 <223> OTHER INFORMATION: a factor Xa cleavage site  
 132 <400> SEQUENCE: 8  
 133 Leu Val Pro Arg Gly Ser  
 134 1 5  
 137 <210> SEQ ID NO: 9  
 138 <211> LENGTH: 8  
 139 <212> TYPE: PRT

C--> 102 <213> ORGANISM: Artificial  
 W--> 106 <223> OTHER INFORMATION: *exploration mandatory in <2207-2237 section*  
 W--> 106 <400> 6 *(see p. 6 for even exploration)*

## RAW SEQUENCE LISTING

DATE: 08/26/2004

PATENT APPLICATION: US/09/858,332E

TIME: 11:22:51

Input Set : D:\SEQUENCE LISTING.TXT

Output Set: N:\CRF4\08262004\I858332E.raw

```

140 <213> ORGANISM: Artificial Sequence
142 <220> FEATURE:
143 <223> OTHER INFORMATION: a thrombin cleavage site
145 <400> SEQUENCE: 9
146 His Pro Phe His Leu Val Ile His
147 1 5
150 <210> SEQ ID NO: 10
151 <211> LENGTH: 10
152 <212> TYPE: PRT
153 <213> ORGANISM: Artificial Sequence
155 <220> FEATURE:
156 <223> OTHER INFORMATION: a renin cleavage site
158 <400> SEQUENCE: 10
159 Cys Tyr Pro Tyr Asp Val Pro Asp Tyr Ala
160 1 5 10
163 <210> SEQ ID NO: 11
164 <211> LENGTH: 8
165 <212> TYPE: PRT
166 <213> ORGANISM: Artificial Sequence
168 <220> FEATURE:
169 <223> OTHER INFORMATION: an immunological tag
171 <400> SEQUENCE: 11
172 Asp Tyr Lys Asp Asp Asp Asp Lys
173 1 5
176 <210> SEQ ID NO: 12
177 <211> LENGTH: 11
178 <212> TYPE: PRT
179 <213> ORGANISM: Artificial Sequence
181 <220> FEATURE:
182 <223> OTHER INFORMATION: an immunological tag
184 <400> SEQUENCE: 12
185 Cys Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu
186 1 5 10
189 <210> SEQ ID NO: 13
190 <211> LENGTH: 3426
191 <212> TYPE: DNA
192 <213> ORGANISM: Artificial Sequence
194 <220> FEATURE:
195 <223> OTHER INFORMATION: DNA sequence of vector containing cDNA of
196 recombinant enterokinase
198 <400> SEQUENCE: 13
199 gacgaaaggg cctcgtgata cgcctatattt tatagggttaa tgtcatgata ataattggttt 60
200 cttagacgtc aggtggcact ttctggggaa atgtgcgcgg aaccctatt tgtttatattt 120
201 tctaaataca ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat 180
202 aatattgaaa aaggaagagt atgagtattc aacatttcg tgcgcctt attccctttt 240
203 ttgcggcatt ttgccttcct gtttttgctc acccagaaac gctggtgaaa gtaaaagatg 300
204 ctgaagatca gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga 360
205 tccttgagag ttttcgcccc gaagaacgtt ttccaatgat gagcactttt aaagttctgc 420
206 tatgtggcgc ggtattatcc cgtattgacg cggggcaaga gcaactcggc gcgcgcatac 480

```

## RAW SEQUENCE LISTING

DATE: 08/26/2004

PATENT APPLICATION: US/09/858,332E

TIME: 11:22:51

Input Set : D:\SEQUENCE LISTING.TXT

Output Set : N:\CRF4\08262004\I858332E.raw

```

207 actatttctca gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg 540
208 gcatgacagt aagagaatta tgcagtgtcg ccataaccat gagtataaac actgcggcca 600
209 acttacttct gacaacgacg ggaggaccga aggagctaac cgcttttttg cacaacatgg 660
210 gggatcatgt aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg 720
211 acgagcgtga caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg 780
212 gcgaactact tactctagct tccggccaac aattaataga ctggatggag gcggataaag 840
213 ttgcaggacc acttctgcgc tcggcccttc cggttggtcg gtttattgct gataaatctg 900
214 gagccggtga gcgtgggtct cgcggtatca ttgcagcact gggggccagat ggtaagccct 960
215 cccgtatcgt agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac 1020
216 agatcgctga gataggtgcc tctactgatta agcattggta actgtcagac caagtttact 1080
217 catatatact ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga 1140
218 tcctttttga taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagggg 1200
219 cagaccccggt agaaaagatc aaaggatctt cttgagatcc tttttttctg cgcgtaatct 1260
220 gctgcttgca aacaaaaaaa ccaccgtac cagcggtggt ttgtttgccg gatcaagagc 1320
221 tacciaactct ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgtcc 1380
222 ttctagtgtg gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc 1440
223 tcgctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg 1500
224 ggttggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga acgggggggt 1560
225 cgtgcacaca gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg 1620
226 agctatgaga aagcgcacg cttcccgaag ggagaaaggc ggacaggat cgggtaagcg 1680
227 gcagggtcgg aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggatatctt 1740
228 atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgcctcgtcag 1800
229 gggggcggag cctatggaaa aacgccagca acgcggcctt tttacggttc ctggcctttt 1860
230 gctggccttt tgctcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta 1920
231 ttaccgcctt tgagttagct gataccgctc gccgcagccg aacgaccgag cgcagcagat 1980
232 cagttagcga ggaagcggaa gagcgcccaa tacgcaaac gcctctcccc gcgcgttggc 2040
233 cgattcatta atgcagctgg cacgacaggt tccccgactg gaaagcgggc agtgagcgca 2100
234 acgcaattaa tgtgagttag ctactcatt aggcacccca ggctttacac tttatgcttc 2160
235 cggtcgtat gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg 2220
236 accatgatta cgccaagctt gaaggatcat ctcatccaca atgtccaca agaggagcac 2280
237 gctcatgccc acaacaagat cgatattgtc ggaggaaagt actccagaga aggagcctgg 2340
238 ccttgggtcg ttgctctgta ttctgacgat caacaggctc gcggagcttc tctggtgagc 2400
239 agggattggc tgggtgcggc cgcccactgc gtgtacggga gaaatatgga gccgtctaa 2460
240 tggaaaagcag tgctaggcct gcatatggca tcaaatctga cttctctca gatagaaact 2520
241 aggttgattg accaaattgt cataaaccca cactacaata aacggagaaa gaacaatgac 2580
242 attgccatga tgcattctga aatgaaagtg aactacacag attatataca gcctatttgt 2640
243 ttaccagaag aaaatcaagt ttttcccca ggaagaattt gttctattgc tgggtggggg 2700
244 gcacttatat atcaaggttc tactgcagac gtactgcaag aagctgacgt tccccttcta 2760
245 tcaaatgaga aatgtcaaca acagatgcc gaatataaca ttacggaaaa tatggtgtgt 2820
246 gcaggctatg aagcaggagg ggtagattct tgtcaggggg attcaggcgg accactcatg 2880
247 tgccaagaaa acaacagatg gctcctggct ggcgtgacgt catttgata tcaatgtgca 2940
248 ctgcctaata gccaggggt gtatgcccg gtcccaagg tccacagagt gatacaaagt 3000
249 tttctacatg agctcgtaat tagctgagaa ttcactggcc gtcgttttac aacgtcgtga 3060
250 ctgggaaaaa cctggcggtta cccaacttaa tcgccttgca gcacatcccc ctttcgccag 3120
251 ctggcgtaat agcgaagagg ccgcaccga tcgccttcc caacagttgc gcagcctgaa 3180
252 tggcgaatgg cgcctgatgc ggtattttct ccttaogcat ctgtgcggta tttcacaccg 3240
253 catatggtgc actctcagta caatctgctc tgatgccgca tagttaagcc agccccgaca 3300
254 cccgccaaaca cccgctgacg cgccctgacg ggcttgtctg ctcccggcat ccgcttacag 3360
255 acaagctgtg accgtctccg ggagctgcat gtgtcagagg ttttcaccgt catcaccgaa 3420

```

## RAW SEQUENCE LISTING

DATE: 08/26/2004

PATENT APPLICATION: US/09/858,332E

TIME: 11:22:51

Input Set : D:\SEQUENCE LISTING.TXT

Output Set: N:\CRF4\08262004\I858332E.raw

```

256 acgcgc 3426
258 <210> SEQ ID NO: 14
259 <211> LENGTH: 269
260 <212> TYPE: PRT
261 <213> ORGANISM: Artificial Sequence
263 <220> FEATURE:
264 <223> OTHER INFORMATION: protein sequence of vector containing cDNA of
265 recombinant enterokinase
267 <400> SEQUENCE: 14
268 Met Thr Met Ile Thr Pro Ser Leu Lys Asp His Leu Ile His Asn Val
269 1 5 10 15
270 His Lys Glu Glu His Ala His Ala His Asn Lys Ile Asp Ile Val Gly
271 20 25 30
272 Gly Ser Asp Ser Arg Glu Gly Ala Trp Pro Trp Val Val Ala Leu Tyr
273 35 40 45
274 Phe Asp Asp Gln Gln Val Cys Gly Ala Ser Leu Val Ser Arg Asp Trp
275 50 55 60
276 Leu Val Ser Ala Ala His Cys Val Tyr Gly Arg Asn Met Glu Pro Ser
277 65 70 75 80
278 Lys Trp Lys Ala Val Leu Gly Leu His Met Ala Ser Asn Leu Thr Ser
279 85 90 95
280 Pro Gln Ile Glu Thr Arg Leu Ile Asp Gln Ile Val Ile Asn Pro His
281 100 105 110
282 Tyr Asn Lys Arg Arg Lys Asn Asn Asp Ile Ala Met Met His Leu Glu
283 115 120 125
284 Met Lys Val Asn Tyr Thr Asp Tyr Ile Gln Pro Ile Cys Leu Pro Glu
285 130 135 140
286 Glu Asn Gln Val Phe Pro Pro Gly Arg Ile Cys Ser Ile Ala Gly Trp
287 145 150 155 160
288 Gly Ala Leu Ile Tyr Gln Gly Ser Thr Ala Asp Val Leu Gln Glu Ala
289 165 170 175
290 Asp Val Pro Leu Leu Ser Asn Glu Lys Cys Gln Gln Gln Met Pro Glu
291 180 185 190
292 Tyr Asn Ile Thr Glu Asn Met Val Cys Ala Gly Tyr Glu Ala Gly Gly
293 195 200 205
294 Val Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Met Cys Gln Glu
295 210 215 220
296 Asn Asn Arg Trp Leu Leu Ala Gly Val Thr Ser Phe Gly Tyr Gln Cys
297 225 230 235 240
298 Ala Leu Pro Asn Arg Pro Gly Val Tyr Ala Arg Val Pro Arg Phe Thr
299 245 250 255
300 Glu Trp Ile Gln Ser Phe Leu His Glu Leu Val Ile Ser
301 260 265
304 <210> SEQ ID NO: 15
305 <211> LENGTH: 12
306 <212> TYPE: PRT
307 <213> ORGANISM: Artificial Sequence
309 <220> FEATURE:
310 <223> OTHER INFORMATION: an amino acid sequence embodiment of the affinity

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/858,332E

DATE: 08/26/2004  
TIME: 11:22:52

Input Set : D:\SEQUENCE LISTING.TXT  
Output Set: N:\CRF4\08262004\I858332E.raw

*PYI*

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:21; Xaa Pos. 2,3,5,6,7,9

Seq#:23; Xaa Pos. 2,3,5,6,8,9

Use of <220> Feature(NEW RULES):

*error explanation*  
Sequence(s) are missing the <220> Feature and associated headings.

Use of <220> to <223> is MANDATORY if <213> ORGANISM is "Artificial Sequence" or "Unknown". Please explain source of genetic material in <220> to <223> section (See "Federal Register," 6/01/98, Vol. 63, No. 104, pp.29631-32) (Sec.1.823 of new Rules)

Seq#:6

## VERIFICATION SUMMARY

DATE: 08/26/2004

PATENT APPLICATION: US/09/858,332E

TIME: 11:22:52

Input Set : D:\SEQUENCE LISTING.TXT

Output Set: N:\CRF4\08262004\I858332E.raw

L:102 M:220 C: Keyword misspelled or invalid format, &lt;213&gt; ORGANISM for SEQ ID#:6

L:106 M:258 W: Mandatory Feature missing, &lt;223&gt; Tag not found for SEQ#:6, &lt;213&gt;

ORGANISM:Artificial Sequence

L:106 M:258 W: Mandatory Feature missing, &lt;223&gt; Blank for SEQ#:6,Line#:106

L:394 M:258 W: Mandatory Feature missing, &lt;220&gt; Tag not found for SEQ ID#:21

L:399 M:258 W: Mandatory Feature missing, &lt;220&gt; Tag not found for SEQ ID#:21

L:404 M:258 W: Mandatory Feature missing, &lt;220&gt; Tag not found for SEQ ID#:21

L:409 M:258 W: Mandatory Feature missing, &lt;220&gt; Tag not found for SEQ ID#:21

L:414 M:258 W: Mandatory Feature missing, &lt;220&gt; Tag not found for SEQ ID#:21

L:415 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21 after pos.:0

L:445 M:258 W: Mandatory Feature missing, &lt;220&gt; Tag not found for SEQ ID#:23

L:449 M:258 W: Mandatory Feature missing, &lt;220&gt; Tag not found for SEQ ID#:23

L:453 M:258 W: Mandatory Feature missing, &lt;220&gt; Tag not found for SEQ ID#:23

L:457 M:258 W: Mandatory Feature missing, &lt;220&gt; Tag not found for SEQ ID#:23

L:461 M:258 W: Mandatory Feature missing, &lt;220&gt; Tag not found for SEQ ID#:23

L:462 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23 after pos.:0